

Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/105513/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Preece, Alun ORCID: <https://orcid.org/0000-0003-0349-9057>, Spasic, Irena ORCID: <https://orcid.org/0000-0002-8132-3885>, Evans, Kieran ORCID: <https://orcid.org/0000-0003-0414-0812>, Rogers, David, Webberley, William, Roberts, Colin ORCID: <https://orcid.org/0000-0003-0595-2740> and Innes, Martin ORCID: <https://orcid.org/0000-0002-8950-8147> 2018. Sentinel: a co-designed platform for semantic enrichment of social media streams. IEEE Transactions on Computational Social Systems 5 (1), pp. 118-131. 10.1109/TCSS.2017.2763684 file

Publishers page: <http://dx.doi.org/10.1109/TCSS.2017.2763684>
<<http://dx.doi.org/10.1109/TCSS.2017.2763684>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Sentinel: A Codesigned Platform for Semantic Enrichment of Social Media Streams

Alun Preece¹, Irena Spasić, Kieran Evans, David Rogers, William Webberley, Colin Roberts, and Martin Innes

Abstract—We introduce the Sentinel platform that supports semantic enrichment of streamed social media data for the purposes of situational understanding. The platform is the result of a codesign effort between computing and social scientists, iteratively developed through a series of pilot studies. The platform is founded upon a knowledge-based approach, in which input streams (*channels*) are characterized by spatial and terminological parameters, collected media is preprocessed to identify significant terms (*signals*), and data are tagged (*framed*) in relation to an ontology. Interpretation of processed media is framed in terms of the 5W framework (who, what, when, where, and why). The platform is designed to be open to the incorporation of new processing modules, building on the knowledge-based elements (channels, signals, and framing ontology) and accessible via a set of user-facing apps. We present the conceptual architecture for the platform, discuss the design and implementation challenges of the underlying stream-processing system, and present a number of apps developed in the context of the pilot studies, highlighting the strengths and importance of the codesign approach and indicating promising areas for future research.

Index Terms—Natural language processing, social computing, streaming media, user centered design.

I. INTRODUCTION

IT HAS become widely recognized that public social media streams can provide valuable insight and actionable information to support situational understanding and decision making [1]–[3]. Twitter, in particular, has emerged as a rich source of publicly accessible data that can reveal and track events, issues, and trends in real time. The low unit cost of generating and consuming tweets (restricted to 140 characters until late 2017) has made them useful for rapid information dissemination, e.g., of breaking news and eyewitness reports, and mobilizing individuals into collective behavior, e.g., protesting an issue or campaigning for support [4], [5]. Other forms of public social media, e.g., blog posts or comment threads on news articles, provide complementary platforms for raising issues, sharing information, and discussion.

While it is generally acknowledged that relying on social media to provide a balanced understanding of a situation is

highly risky, not least for demographic reasons [6] and in terms of information quality [7], this does not diminish its value as a source of insight that can and should be combined with other sources to achieve situational understanding. In this sense, social media serves as a sensor, providing data that can be combined with findings from other sources to build a more complete picture as in traditional information fusion and sensemaking approaches [8], [9]. Data gleaned from social media streams like Twitter can be valuable both in real-time and post hoc analyses.

Recognizing this value in social media, a variety of analysis tools and platforms have been produced in recent years, in both the commercial and academic sectors. Many of the commercial tools have a focus on marketing and brand-management applications; examples include Blurr (www.blurr.co.uk), Hootsuite (hootsuite.com), Social Studio (www.marketingcloud.com), and RepKnight (www.repknights.com). These tools tend to take a black box approach, supporting a number of specific analyses but being hard to repurpose for new applications, or to integrate new functionality.

This paper grew out of a collaboration between social scientists focused on the practice and science of policing and computer scientists with research interests in data mining and decision support. Our principles for the work were as follows.

- 1) To create a glass box platform for semantic enrichment of social media data to support situational understanding, designed to be as open as possible to the integration of new components, models, data sources, and user interfaces. By *semantic enrichment*, we mean the automatic and semiautomatic integration of metadata, defined by a semantic data model, into data products derived from social media.
- 2) To ensure that the platform is well fitted to the needs of its end users by adopting a codesign approach, with social scientists and subject-matter experts closely involved throughout. By *codesign*, we mean that end users are encouraged and supported in designing tools for themselves, with computer scientists and software engineers acting as facilitators and implementers.
- 3) To use knowledge technologies as a foundation for the platform, so that processes of semantic enrichment are embedded throughout the components and information flows between them. By *knowledge technologies*, we mean techniques derived from the fields of knowledge representation and reasoning, natural language processing (NLP), and data mining.

A sizeable amount of research and development has been done in this domain, highlighting the need for meaningful data

Manuscript received September 7, 2017; accepted September 28, 2017. Date of publication December 20, 2017; date of current version February 23, 2018. This work was supported in part by the European Commission through the Project Tackling Radicalisation in Dispersed Societies and in part by the Economic and Social Research Council through the Project After Woolwich: Social Reactions on Social Media under Grant ES/L008181/1, and in part by Cardiff University. (Corresponding author: Alun Preece.)

The authors are with the Crime and Security Research Institute, Cardiff University, Cardiff CF10 3AE, U.K. (e-mail: preecead@cardiff.ac.uk; spasic@cardiff.ac.uk; evansk12@cardiff.ac.uk; rogersdm1@cardiff.ac.uk; webberleywm@cardiff.ac.uk; robertsc10@cardiff.ac.uk; innesm@cardiff.ac.uk).

Digital Object Identifier 10.1109/TCSS.2017.2763684

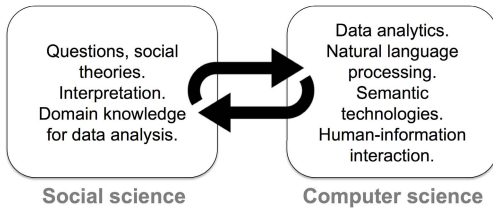


Fig. 1. Interdisciplinary social/computer science codesign approach.

visualization, interactivity with users, and scalable processing among other key features; for recent surveys, see [10]–[12]. In addition to these areas, we emphasize the importance of open approaches, with end users (analysts) driving the design process. In this paper, the underlying requirement for openness and codesign is the need to incorporate bespoke models into the analysis system, representing elements of social science theories.

Fig. 1 shows what we viewed as the synergies between the social science and computer science teams, that the codesign approach was intended to exploit. In the early stages of the work, hypotheses and theories tended to originate from the social scientists while the computer scientists, being schooled in what was technologically possible, generated ideas for data analyses, semantic enrichment (to reveal meaning), and visualization. The social science team provided background knowledge to support the data mining and enrichment and interpretation of results, typically leading to formulation of further questions and hypotheses. Over time, roles became less well defined, e.g., with members of the computer science team originating questions or hypotheses, and social scientists suggesting analysis rules or outline algorithms.

Development of our platform, named Sentinel (“Semantic Intelligence”), was framed by a number of pilot studies drawn from ongoing social science work in policing the following.

- 1) Scanning of social media traffic in relation to geographic regions, including a major city and a medium-density city region.
- 2) A longitudinal study of a high-profile crime and its effects over a ten-month period from perpetration to sentencing.
- 3) A real-time study of a major planned event in a city region, including the buildup over a three-month period.

The three studies selected for this paper exemplify our particular requirements for openness and codesign, including the incorporation of bespoke social science models, as well as the more general requirements to provide meaningful, interactive data visualizations and scalable processing of social media data considered as a sensor stream.

This paper is organized as follows. Section II introduces the conceptual architecture of the Sentinel platform, focusing on the key design choices and semantic models. Section III describes the data analysis architecture in more detail, with particular emphasis on the pipeline designed to support semantic enrichment of real-time streamed data. Section IV presents some of the user-facing Sentinel applications (apps) developed to date, highlighting features incorporated as a result of codesign. Section V draws on our experience using the

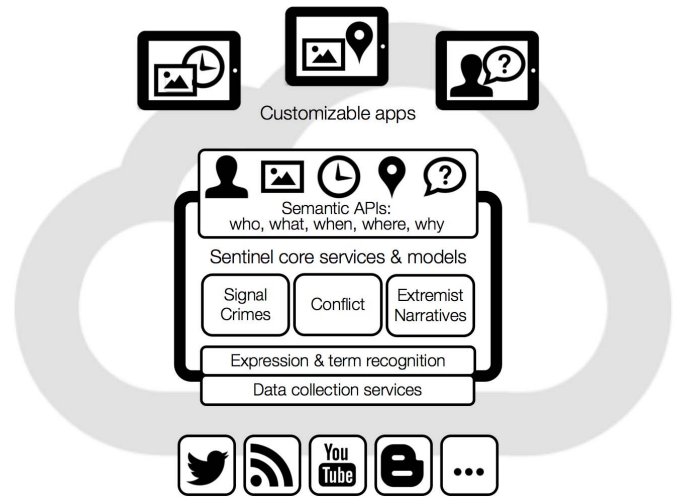


Fig. 2. Sentinel platform conceptual architecture.

platform in the pilot studies, and evaluates the extent to which Sentinel meets the objectives of supporting interdisciplinary social/computing science in an open and flexible manner. Section VI places Sentinel in the context of related work, and Section VII offers concluding discussion and future work.

II. PLATFORM ARCHITECTURE

An overview of the conceptual architecture of the Sentinel platform is shown in Fig. 2. The core platform comprises a set of cloud services, which process data collected from a variety of social media streams illustrated at the bottom of the figure, and deliver interpreted information to end users via a set of apps shown at the top. The platform is designed to be open at the top (to the rapid creation of new apps), the bottom (to the incorporation of new social media feeds), and the middle (to extension with new data analysis and modeling services).

Given the team’s interest in areas relating to policing and society, originally our interest centered on four main social media streams: 1) Twitter, providing indication of real-time events and issues, information dissemination, and collective behaviors such as campaigning; 2) community blogs, providing richer coverage of issues and events, as well as discussion and information sharing; 3) comments on YouTube videos, highlighting (via text) features of the videos, and reaction to them; and 4) readers’ comments on news reports on mainstream media websites. In accordance with the terms-of-use for the respective services, data collection services were created for each of the four kinds of media. Subsequent work, however, has focused on Twitter as the primary driver for Sentinel, since tweets are frequently used as link carriers to other social media, e.g., a tweeted link to a news story, a YouTube video, or a blog posting. This paper will, therefore, focus mainly on how Sentinel uses Twitter data.

Data collection in Sentinel is organized and managed by means of semantic *channels*. A channel is associated with one or more social media feeds and provides a bridge between *what* the user is interested in (in broad or narrow terms, e.g., social media in a particular geographic region or

TABLE I
SENTINEL CHANNEL PARAMETERS AND THEIR APPLICATION TO TWITTER DATA COLLECTION

Channel parameter	Collected data (Twitter API)
at most one <i>geo bounding box</i>	geotagged tweets within the box
zero or more <i>spatial terms</i> (names of places)	tweets matching any specified place name
zero or more <i>topic terms</i> (words, phrases or tags)	tweets matching any specified topic term
zero or more <i>names of actors</i> (real names or social media names)	tweets matching any specified actor name
zero or more <i>social media feeds</i> (specific social media accounts)	tweets from any specified account

relating to a set of topics) and *how* to collect relevant data (e.g., tweets geotagged as originating within a particular geospatial bounding box, or matching a set of Twitter search terms). Collection channels are set to run in real time for whatever period the user desires. Channel parameters are shown in Table I; when applied to Twitter, these parameters specify an endpoint for the streaming application program interface (API) (dev.twitter.com/streaming). The primary purpose of a channel is to specify information requirements as a first step in information retrieval, when faced with an enormous potential volume of available social media data. As with any internet search, the choice of parameters used to express the information need is something of an art, and in practice, channel parameters tend to undergo refinement using feedback over the course of a particular project. Sentinel is designed to allow easy creation and modification of channels. Further discussion of practical experience with channels in relation to the pilot studies appears in Section V.

Since all of our social media sources are text-based (even though some contain links to images or video), Sentinel initially applies a set of low-level NLP services to them, applying a number of filters and performing automatic term recognition (see Section III) in an attempt to identify the main topics in the current conversation. These terms emerge from the collected data in a bottom-up fashion, so we refer to them as *signals*.

Higher level NLP services are then applied to media in relation to detected signals, ranging from simple sentiment and social-network analyses to more complex models drawn from the team's social science research (see Section IV). At the higher levels of the core Sentinel architecture, data interpretation is framed by the 5W model using an ontology described in the following. The concepts in the ontology, which is intended to be highly extensible, are used to provide a semantic *framing* of signals and processed media. A collection of semantic APIs allows processed data to be accessed by end users via the apps.

A. Bottom-Up Data Interpretation: Signaling

Generation of signal terms is performed in Sentinel using the FlexiTerm automatic term recognition (ATR) algorithm [13]. In principle, other ATR approaches could be plugged-in; however, we designed FlexiTerm specifically to operate robustly on social media data (originally blogs). FlexiTerm initially performs linguistic filtering to select term candidates, which are noun phrases. It then ranks candidates based on their termhood, a measure calculated as a combination of frequency and collocational stability. In order to

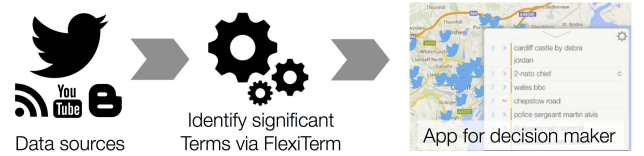


Fig. 3. Bottom-up (data-driven) interpretation of social media streams.

improve the quality of termhood calculation, which may be affected by term variations, FlexiTerm uses a range of methods, including managing syntactic variation (e.g., “English Defence League’s (EDL) leader” versus “leader of the EDL”) by using a bag-of-words approach, and handling orthographic (e.g., “Anglo-Saxon” versus “anglo saxon”) and morphological (e.g., “England” versus “English”) variations using stemming in combination with lexical and phonetic similarity measures. The latter helps correct for common writing styles on social media, including variable spelling (e.g., “Lee” versus “Leigh”), misspelling (e.g., “Woolwich” versus “Woollich”), and abbreviations (e.g., “2nite” versus “tonight”).

This data-driven (bottom-up) approach to extracting meaning from streams of social media to inform an end user is shown in Fig. 3. Commonly, the signal terms are used as a first stage in further processing, even though under some circumstances they can be informative as a direct output of Sentinel, particularly when users are monitoring a situation in real time. Being noun phrases, the signal terms are often better-contextualized, and, therefore, more informative than Twitter’s trending topics (e.g., the term “armed police” rather than the words “armed” or “police”) as well as being generally more relevant to a particular area of interest due to Sentinel’s channeling approach. Sentinel also allows detected terms to be added to the ontology and used in top-down interpretation.

B. Top-Down Data Interpretation: Framing

An ontology is defined as an explicit conceptualization of a domain through a set of concepts, their definitions, and relations between them [14]. The purpose of an ontology is to provide effective means of communication within a domain. This communication can be between humans, between computer systems, or between humans and computer systems [15]. Each ontology is characterized by its domain, purpose, and formality.

The use of ontologies in processing social media data is reasonably well established (see [10] for a recent survey); however, in view of our objective to support interpretation of

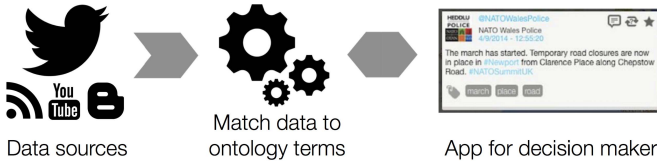


Fig. 4. Top-down (ontology-driven) interpretation of social media streams.

these data in terms of bespoke social science models, our aim was not to reuse one or more existing ontologies but rather to support the use of custom models for user-led studies. To this end, we initially created an ontology of concepts relating to the domain of policing and society. Our *Sentinel ontology* defines 479 concepts. Its purpose is twofold. First, it supports consistent and unambiguous knowledge sharing between the team members with different specialisms and as such represents the foundation of our codesign approach. Second, the large volumes of Twitter data make it difficult to efficiently locate, retrieve, and manage actionable information without the use of text mining applications. In order to interpret incoming data streams efficiently, text evidence needs to be linked to the ontology as the main repository of formally represented domain knowledge. In order to support text mining applications within the Sentinel framework, the ontology includes 389 synonyms in addition to preferred concept names. The ontology is specifically intended to help frame incoming data streams by automatically tagging the text content with ontology concepts. This allows an end user to browse or query collected media in terms of the domain semantics formally modeled by ontology, e.g., by looking for references to a particular crime or group. In this sense, the ontology supports data interpretation in terms of who, what, and where questions from the 5W model. We view this process as top-down in terms of how the user accesses the data; however, the flow of data through the collection channels is unchanged, and ontological tagging of social media is performed continuously as media are collected. This process is shown in Fig. 4: the bidirectional arrow between user and processing services is intended to show ontology-based queries originating with the user while, in Fig. 3, they were a passive receiver of signals.

In addition to concept definitions and the vocabulary of their names, the *is_a* relationship provides a taxonomic structure of the ontology, with maximum depth of eight nodes and four child nodes per concept on average. We used the basic formal ontology (BFO) as the upper level ontology [16]: the top-most class is *entity*, which is divided into *continuant* and *occurrent* (see Fig. 5). The next level concepts reused from BFO are *material entity*, *realizable entity*, *quality*, and *process*. Demonstrating the incorporation of bespoke social science models, the lower level concepts represent our team's own framework of crime and social disorder based on the signal crimes theory [17] and drawing on extensive fieldwork in community and neighborhood policing [18]. In addition to the aforementioned hierarchical organization of concepts, a network of six other named relationships (e.g., *part_of*, *associated_with*, and so on) with a total of 132 relationship instances allows a user to: 1) navigate through large amounts

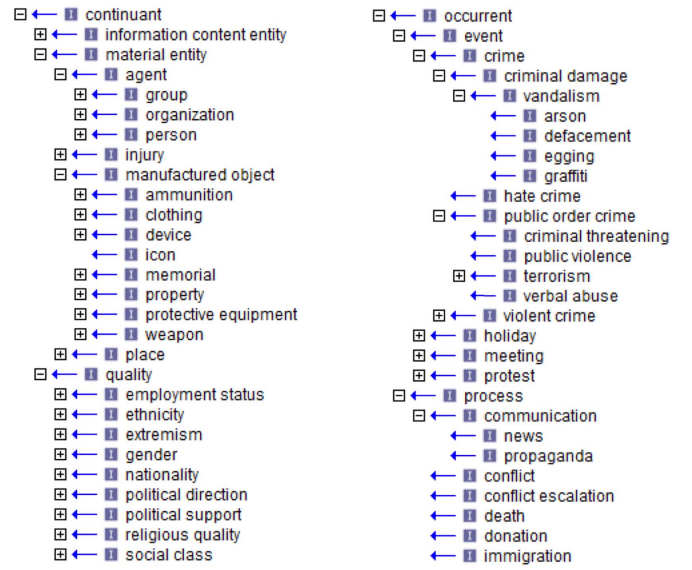


Fig. 5. Extract of Sentinel ontology concepts related to policing and society.

```
[Term]
id: SENTINEL:0000089
name: hate crime
def: "Hate crime occur when a perpetrator targets a victim because of his or her perceived membership in a certain social group. Examples of such groups include but are not limited to: racial group, religion, sexual orientation, ethnicity or gender identity." [wikipedia:Hate_crime]
synonym: "bias-motivated crime" EXACT []
synonym: "race hate" EXACT []
is_a: SENTINEL:0000007 ! crime
relationship: targets SENTINEL:0000138 ! ethnicity
relationship: targets SENTINEL:0000132 ! religious group
relationship: targets SENTINEL:0000137 ! nationality
```

Fig. 6. Example concept description in the Sentinel ontology.

of ontology-tagged data by following a hierarchy of related terms (e.g., a retrieved set of tweets tagged with “arson” can be expanded to include all types of “vandalism” and vice versa) and 2) perform an implicit search, i.e., to access implicitly stated relevant information without having to explicitly name all relevant information (e.g., searching for “terrorism,” a concept *associated_with* a “proscribed organization,” will also retrieve mentions of its subconcepts, e.g., “Al-Qaeda”).

Finally, the Sentinel ontology is formally encoded in OBO flat file format. OBO is a text file format used by OBO-Edit, an open-source, platform-independent application for viewing and editing ontologies [19]. OBO was chosen over the Web Ontology Language, a Semantic Web language supported by the World Wide Web Consortium [20], due to its simplicity, which is one of the key reasons for its widespread use within the biomedical community. OBO follows the tag-value format, where each tag-value pair consists of a tag name and the tag value (see Fig. 6 for an excerpt from the Sentinel ontology). Such simple human-readable syntax makes it appropriate for collaboration with social scientists who are expected to share a similar level of computational literacy with biomedical experts.

III. STREAM-PROCESSING ARCHITECTURE

Fig. 7 shows an overview of the design of the social media stream-processing pipeline in Sentinel, from data collection to data delivery via user-facing apps. The pipeline is broken down into three phases; the *collection* phase of

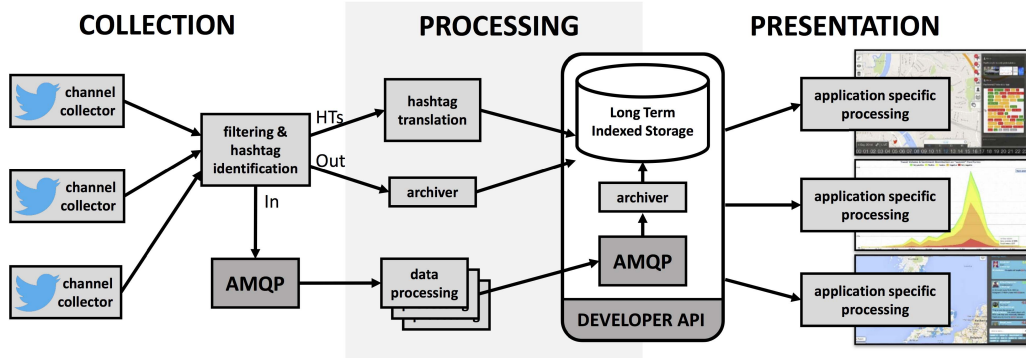


Fig. 7. Sentinel stream-processing architecture overview.

the figure corresponds to the *data collection services* element of Fig. 2, the *processing* phase corresponds to the *Sentinel core services and models* element of Fig. 2, with the *presentation* phase reflecting the numerous applications that have been built upon the data produced. The *API* in both figures separates the core processing from the apps. Given the requirement for Sentinel to support social media stream processing, the architecture is based on Advanced Message Queuing Protocol (AMQP) [21], currently implemented via open-source RabbitMQ (www.rabbitmq.com). AMQP, as a message-oriented middleware standard, provides great flexibility in managing reliable media flow and asynchronous processing through the system. It has also allowed us to keep Sentinel running mostly live since December 2013, because it allows collected media to be queued while changes are made to the pipeline components and configuration.

The discussion here is focused on Twitter data collection; as mentioned in Section II, due to its properties in terms of real-time coverage and operating as a link carrier for other media, development of Sentinel has focused on Twitter as the primary type of social media feed. Collectors for other media have been developed and operate similar to the Twitter mechanisms.

Each data collector is coupled to a filter, which excludes social media posts containing particular terms from being processed further, in order to reduce noise that may skew data analysis; in the case of Twitter, this mechanism is used to remove significant volumes of tweets with common phrases such as “Happy birthday” or the names of celebrities which otherwise tend to dominate channels, especially ones with a significant proportion of geospatially relevant tweets. These excluded tweets are not discarded, but are archived and can be processed at some later date if required by apps.

At this stage, we also extract hashtags and usernames (prefixed with # and @, respectively) from tweets, which are passed to a hashtag translator. Hashtags can often be formed from a number of words concatenated together (e.g., “photo of the day” becomes “#photooftheday”) and so the hashtag translator provides a heuristic approach to decompressing these social media specific lexical features. The translator works using a series of regular expressions designed to split hashtags into a set of words based on a series of common practices, such as camel casing (“#PhotoOfTheDay”), underscoring

(“#photo_of_the_day”), and hyphenation (“#photo-of-the-day”). The translation with the highest number of known words is then added to a database table along with its score, and with a fully lowercased version of the hashtag being used as its identifier. This database is periodically scanned in order to identify the translation for each identifier with the highest word score, which is then cached in a second table that acts as a lookup table that is available in the API for middle and end applications. This approach allows for the most appropriate translation to emerge over time.

Multiple collectors run concurrently, hosted on a cloud (implemented on open-source OpenStack, www.openstack.org), and pass the data to the first AMQP message queue. The higher level processing components of Sentinel, including the FlexiTerm ATR module, then draw messages from the queue for further processing as described in Section II. Again, multiple processing modules run concurrently, including multiple instances of the same module where necessary for performance reasons. The higher level processing includes not only signal generation via FlexiTerm but also sentiment analysis, tagging with ontology concepts to frame tweets semantically, and other forms of processing used by the apps (described in Section IV). Associating ontology terms with the tweets is done by the Ontology Indexer module (shown as one of the data processing modules in Fig. 7), which currently performs soft string matching (making it robust to high incidence of typos in social media data) on tweets against the Sentinel ontology (secondstring.sourceforge.net). All semantic products of the higher level processing—including signals, frames, sentiment scores, and others—are stored in the database shown, for use by the apps via the API. The API also gives Sentinel apps access to the media stream via a second AMQP queue. The database is currently implemented using the MongoDB open-source document database (www.mongodb.org), though we anticipate that the next version of Sentinel will use PostgreSQL (www.postgresql.org).

Stability and performance have been key considerations throughout the development of Sentinel. As indicated in Section V, stable operation of Sentinel has been critical to a number of social science studies, some involving real-time situational understanding. While the design of the pipeline was

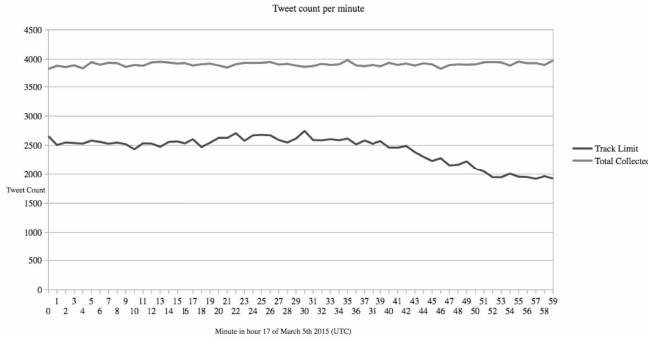


Fig. 8. Collected tweets and “track limit” rates for global geotagged tweets over 1 h.

informed by team members’ previous experience in cloud-based workflow system research [22], the overriding principle for Sentinel was to keep the system as simple and flexible as possible, hence the choice of AMQP as a standard, and reliance on widely used open-source software rather than bespoke components.

In addition to monitoring and tuning the performance of the system while operational as part of the pilot studies, we performed benchmarking to “stress test” the pipeline. The Twitter streaming API is limited to serving approximately 1% of the total volume of tweets generated in real time. With only a few occasional exceptions, mentioned in Section V, none of our channels have regularly exceeded this volume. To test the upper bound of available data, we configured a channel to collect geotagged tweets at a global level. Fig. 8 shows that the global collection received notification from the streaming API that the number of Tweets served up had been rate limited. This demonstrates that the Sentinel pipeline is capable of dealing with the maximum volume of tweets available from Twitter via the streaming API. The global collection was deployed on eight “m2.large” instances on our Openstack cloud. Each “m2.large” instance comprised of 4-V CPUs, 16-GB RAM, and 160 GB of disk space.

IV. USER INTERFACES

We introduce three of the user-facing apps that have been implemented via rapid application development. The development team included three members, two of whom were also involved in the development of the core Sentinel system and its API, while one worked only above the API. We thus endeavored to maintain a separation so that the app developers did not need knowledge of Sentinel below its API. Sections IV-A–C describe: 1) our general-purpose *Sentinel* app; 2) the *SentiSum* app which summarizes event sentiment and impact; and 3) the *SentiNow* app which supports real-time geolocated visualization of event impact.

A. Sentinel App

The Sentinel app was designed to: 1) expose analytic functionality to users in an easy-to-consume way and thus gain feedback as part of the cross-disciplinary interaction (see Fig. 1); 2) serve as a general-purpose real-time situation

awareness tool, particularly when the scope is associated with a particular geographic region; and 3) be available as a demonstration of Sentinel’s capabilities. Originally, this app was designed to run in a web browser for use on screen sizes ranging from tablets to wall-mounted displays and projection screens. More recently, a mobile version was developed.

A sample screenshot from the app is shown in Fig. 9. All sample tweets shown in screenshots in this paper are public. This is a default display for a selected channel. Geotagged tweets are plotted on a scrollable and zoomable map (implemented by the Google maps API). A 24-h timeline appears at the bottom of the screen, along with a date selector. Data can be viewed in live mode or previous dates can be selected. Each hourly marker up to the current hour can be selected to show the signals generated by the FlexiTerm algorithm for that hour (signals for the current hour are updated every 5 min). The screenshot shows the signals for 12 noon. Selecting a signal, by clicking on the corresponding term in the pop-up menu, causes the associated tweets to be displayed in the right-hand drawer, and any of these that are geotagged are displayed in red on the map (the default color is blue). In this example, the noun phrase “chepstow road” (a local place) has been identified as a signal by FlexiTerm, because several tweets in the given period mention it. Some of the geotagged tweets appearing in red are actually located on Chepstow Road and constitute eyewitness reports.

Shown below each individual tweet is the ontology-based tags that match the tweet. Some of the terms in this example are quite generic, e.g., the spatial terms *road* and *place*, while others are more significant as they relate to a kind of event: *protest* and *march*. Icons to the top-right of each tweet can be used to locate that tweet on the map (if geotagged), access images linked to the tweet, and discover further information about the tweet and tweeter (e.g., who has retweeted this tweet, and features of the tweeter’s social network). The seven tabs to the left of the drawer allow access to selections of all the tweets for the selected hour on the timeline, including all tweets, all geotagged tweets, all tweets with images (presented as an image collage), and tweets that the app user has favoured. A search facility is also accessible from the right-hand drawer.

In addition to the default map-based view, a timeline view shows the volumes of tweets for the channel and selected signal terms. This is essentially a subset of the view offered by the SentiSum app, as described in the following. Colored bars to the left of the term names in the signal drawer (bottom center) indicate the aggregate sentiment for the text of tweets associated with that signal in the current hour (or 5 min in live mode). These are computed using the Stanford NLP sentiment algorithm [23] and map a scale from *very negative* to *very positive* onto a color spectrum from red to green. Numbers to the left of the signal drawer show term frequencies (showing that Sentinel is able to detect small signals of only a few tweets) and markers show whether the signal is rising in frequency, falling, nonmoving, or new compared to the previous hour.

The majority of features offered by this app were suggested by our end-user social scientists and incorporated into the app design in collaboration with the computer science team

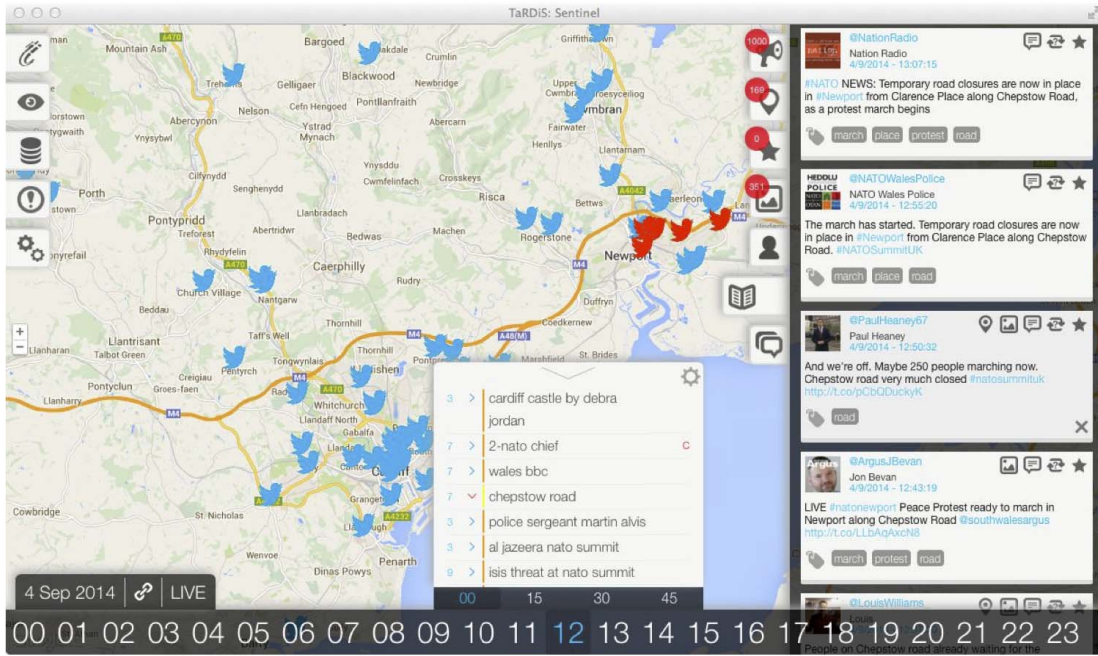


Fig. 9. Sentinel app screenshot.

members. Particular features strongly desired by the end users included the timeline/map choice of display, the easy access to an image collage for a chosen signal term, and the ability to analyze the social network of a tweeter or tweet.

B. SentiSum App

The purpose of the SentiSum app was to support situational understanding in relation to timelines, including identification of key events and their impact in terms of social media traffic, and trends in sentiment. Essentially, it was designed and built rapidly to complement the main Sentinel app, expanding on the functionality of the timeline view but, unlike the original app, showing day-by-day instead of hour-by-hour trends. The app provides two main displays. The first view (see Fig. 10) shows tweet volumes (y-axis) over time (dates on the x-axis), where the volume on each day is divided into five colored bands corresponding to the five high-level sentiment classifications computed by the Stanford algorithm (see the previous section): red/very negative, amber/negative, yellow/neutral, pale green/positive, and dark green/very positive. The second view ignores the volumes and shows only the sentiment profile for tweets on each day (see Fig. 15). The app allows the user to see timeline volumes and profiles for tweets associated with a particular signal term or collection of signal terms; e.g., the user could select a view that shows volumes of tweets and sentiment profile for all tweets associated with signals on a particular channel, including the words “police” and “crime.”

C. SentiNow App

Like SentiSum, the SentiNow app was designed to complement an aspect of the main Sentinel app, in this case the live mode. The display provided by SentiNow is similar in design (but simpler) than the original view shown in Fig. 9

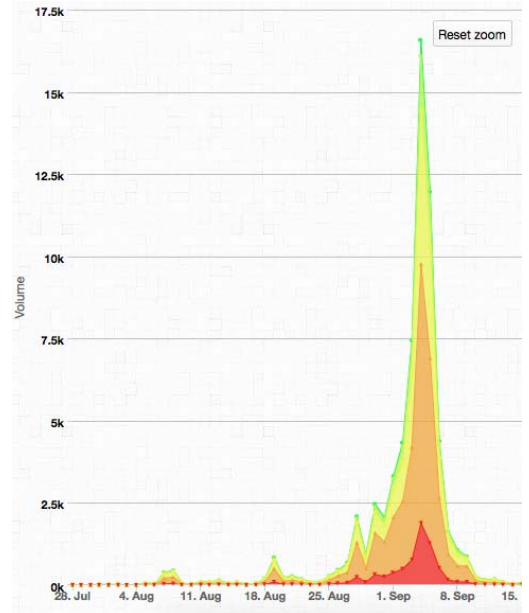


Fig. 10. Volume and sentiment profile view from SentiSum app.

and shows tweets arriving in real time on a map, highlighting the ones matching particular search terms, and showing the sentiment of the matching tweets by varying the color of the Twitter icon on the map from red (very negative) to green (very positive). SentiNow provides a real-time view of the impact of events in terms of where people are talking about those events. SentiNow was built rapidly to provide a “big screen” display for the third pilot study described in Section V.

V. PILOT STUDIES

We present three pilot experiments using the Sentinel platform for various purposes in the context of situational

understanding in relation to policing. These studies were exploratory, conducted to gain experience with the Sentinel platform and identify improvements, while delivering useful results to the social science team members in their studies.

A. Study 1: Regional Situation Awareness

1) *Purpose*: Sentinel was used to monitor Twitter traffic in relation to two geographic regions: 1) a sizeable part of a major city (South London, U.K.) and 2) a medium-sized city region (South Wales, U.K., comprising one city, Cardiff, and a number of smaller towns). Our collaborators were interested particularly in social media relating to public services, including, but not limited to, policing.

2) *Method*: The original version of the Sentinel app described in Section IV was created in August 2013 to support these studies. At that time, data collection for South London was already underway (see the second experiment, in the following); we created a data collection channel for South Wales in the autumn of 2013. Both channels have been running more-or-less continuously since their inception. At the outset, the channels were defined in terms of a geospatial bounding box, to collect geotagged tweets only. In mid-2014, the channel focused around South Wales was expanded to include a number of spatial terms to improve precision and recall in terms of locally relevant tweets. These terms included names of towns, streets, and landmarks (drawing upon police and crime “hotspot” data) and names of key elements of the transport systems (motorway junctions and railway stations).

The channel was tuned at specific times in order to examine periods of relatively high social disorder with particular focus on public–police engagement; e.g., we introduced terms relating to Halloween and Guy Fawkes Night to study impacts of community–police relationships during the period October 31 to November 5, 2014, which in recent years has been characterized by a rise in antisocial behavior in the U.K.

Owing to the long-term nature of the study, real-time “eyes-on” use of the apps was limited to specific periods when there was some expectation of activity, and episodic use to sample the performance of the tools. Multiple users were involved, including analysts from the social science team and policing partners, and testers from the computer science team.

3) *Results*: Sample data volumes from the London and South Wales channels are shown in Fig. 11. The Sentinel architecture was able to handle the data volumes with no performance issues, running on an OpenStack cloud comprised of five Nodes with 160 CPU cores, 34 TB of storage, and 326 GB of RAM. One of the key findings was that, in terms of public services, the dominant issues on both channels tend to be travel-related. Partly, this seems to be because traveling users are more prone to geotag their locations, causing Sentinel to collect their tweets within the channels’ respective bounding boxes, and partly it seems to be that venting one’s travel frustrations is a common usage of Twitter. As an example, Fig. 12 shows part of a screenshot from the Sentinel app applied to the South London channel. Many of the locations of geotagged tweets follow the main railway lines; on a typical morning,

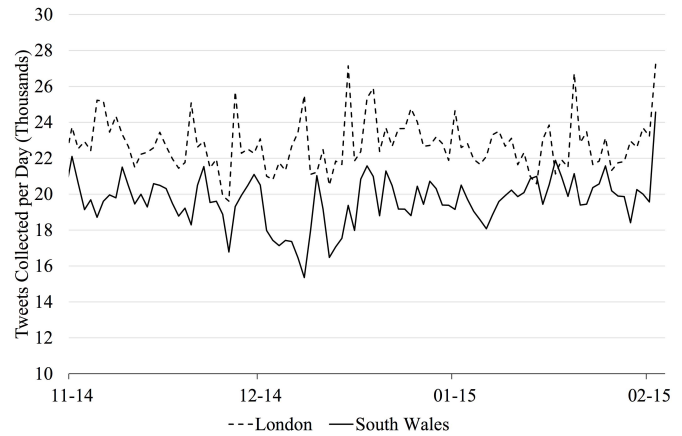


Fig. 11. Sample tweet volumes from London and South Wales channels.

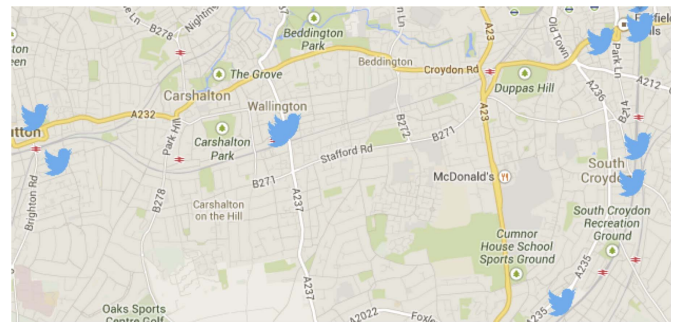


Fig. 12. Geotagged tweets in London corresponding to routes of railway lines.

signals obtained by our FlexiTerm algorithm from the locally generated tweets tend to feature the names of train companies operating in the region: “sw trains” (Southwest Trains) and “southern trains” (Southern Trains).

Social media traffic on the two regional channels also tends to be dominated by major events, including sporting fixtures and public protests. In the summer of 2014, we made minor extensions to the South Wales channel to focus specifically on public protests, where there was potential for social disorder. Places names such as “cardiff” and “newport” were conjoined with event-specific terms such as “protest” and “march” from the ontology. As a result, the channel was able to effectively track unfolding events in terms of generated signals, including minor incidents reported by a relatively small number of tweeters, and campaigning behavior involving retweeting of reports related to the protests.

A striking example of this was a protest march involving over 1500 people in the center of Cardiff on July 26, 2014. At 15:15 one person tweeted: “Watching drunken stag do idiots disrupt peaceful Free Palestine march, Mill Lane Cardiff.” Three mentions of the location “Mill Lane” in the 15:00–16:00 period caused this to be signaled by FlexiTerm (along with several other relevant terms, including “protest in cardiff,” “massive protest in cardiff,” and “cardiff palestine demo”). The other of the “Mill Lane” tweets read: “Cardiff protest turns ugly as it ventures past Walkabout Cardiff and Mill Lane. #cardiff #protest.” No posts were issued on the official South Wales Police account, @swpolice,

around this time, although one of the tweets associated with the signal term “protest in Cardiff” was directed at the official account: “A mass #FreePalestine protest in Cardiff city center and where were the @swpolice”?

By 11 P.M. on July 26, the second-top signal term on the Sentinel app was “riot in mill street cardiff,” triggered by five retweets of a link to a YouTube video of the violence. The next day, the story gathered more attention, with major news outlets picking it up and blaming “poor policing.”¹ By 15:00–16:00 on the 27th, 24 h after the original incident, 46 retweets of the media’s coverage of the violence and police response (“Gaza march violence policing “poor”: Police are criticized over their handling of a protest march in Cardiff a ...)” were picked up by the signal term “protest march in cardiff a” (the “a” at the end of this term is an artifact of the retweets truncating the original quote).

These examples illustrate how Sentinel is able to perform effectively in a bottom-up manner, delivering informative signals generated from relatively small numbers of tweets. The top-down framing of events worked well also: e.g., the tweets in the 15:00–16:00 set on the 27th were tagged with the ontology concepts *march*, *police force*, and *protest*, while the “mill lane” set was also related to concepts *march* and *protest*. The detection of physical-world events such as the violent incident and online events such as the retweeting of the YouTube videos can be of significant importance and value to organizations providing services to the public, including the police and local government. This paper also shows the potential of the Sentinel tools to generate data relevant to understanding public perception of how police resources are deployed with respect to issues seen as causing social harm, including signal crimes [18].

4) *Discussion*: As has been acknowledged elsewhere, accurately locating tweeters is a hard problem [24]. Our region-specific channels are of course accurate in terms of collecting relevant geospatially tagged tweets, as well as posts by Twitter users who provide location data in their profiles. Beyond this, the channels currently rely on mentions of places in the gazetteer part of the channel parameter set. Consequently, these channels provide somewhat limited coverage of social media posts within the target region, and therefore tend to be most effective in relation to: 1) events (large and small) where tweeters tend to mention locations and/or geotag their posts and 2) campaigns, where tweeters again tend to reference places explicitly when seeking local support for some issue (or are retweeting mainstream media posts which again tend to mention locations in the region).

B. Study 2: Tracking the Effects of a High-Profile Crime

1) *Purpose*: Sentinel was used to support a longitudinal study of a high-profile crime and its effects over a ten-month period from perpetration to sentencing. This study began with the murder of Lee Rigby in Woolwich, South London, in May 2013. At that time, our South London

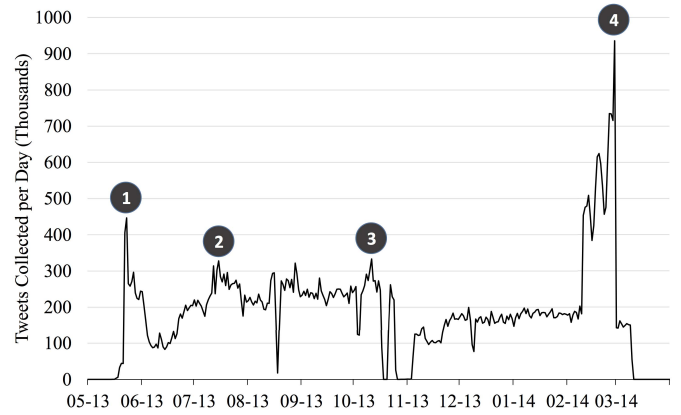


Fig. 13. Tweet volumes from Woolwich channel (key events labeled).

channel described above was already in operation and, therefore, captured a large volume of tweets in relation to the crime, including its initial eye-witness accounts on Twitter. Following the initial event, a separate channel was then created to focus on the incident and its after-effects. From a social science point of view, this channel was designed to support a case study of social reactions to high-profile crimes, to help understand how the general public interprets and make sense of such events. From a computer science viewpoint, the case study provided an opportunity to stress-test the Sentinel architecture over a sustained period. The channel ultimately ran for ten months until after the sentencing of the perpetrators.

2) *Method*: To focus on the specific incident, a thematic channel was created with terms relating to the Woolwich murder, key locations, and names of individuals and groups involved. Over the ten-month period, a number of analysts from the social science fine-tuned the channel in the light of unfolding events. The crime inflamed community tensions and led to protests and a number of criminal incidents, including arson attacks. Rapid modifications were made to the channel’s set of topic terms to capture social media traffic in relation to the evolving situation. As with the previous experiment, there was intermittent “eyes-on” use of the Sentinel app during specific times, most notably in the immediate aftermath of reactionary crimes and disorder, and during the conclusion of the trial and sentencing in February 2014.

3) *Results*: Data volumes from the channel over the ten-month period are shown in Fig. 13. Several key events are labeled: 1) the killing of Lee Rigby (May 22, 2013); 2) Lee Rigby’s funeral (July 12) followed by a large EDL march (July 15); 3) resignation of the leader of the EDL, Tommy Robinson (December 10); 4) sentencing of the murder suspects (February 26). The chart also shows some anomalies due to collection methods, including a number of periods where no data were collected due to system downtime and higher traffic toward the end when data were obtained via Twitter’s commercial API. As in the first study above, occasional system downtime notwithstanding, the Sentinel architecture was able to handle data collection with no performance issues. Findings from the analysis of the tranche of tweets gathered in the first 24 h of the incident highlighted policing issues posed by such highly public crimes regarding the

¹For example, from the BBC: <http://www.bbc.co.uk/news/uk-wales-south-east-wales-28509791>.

permeability of crime scenes. In terms of the leakage of information in the era of social media, the police are no longer able to effectively seal a crime scene where key eyewitnesses are live-tweeting events [25]. Similar observations have been made on relation to other crimes such as the Boston Marathon bombings in 2013.

Analysis of this data set, which totaled over 35 M tweets at the end of the collection period, is still ongoing. Social science findings to date have been reported in [25]–[27]. The aftermath of the crime, social media reports appeared across the U.K. concerning hate crimes targeted toward individuals and religious buildings. Countering the general picture of negative sentiment, one event soon after the murder was striking for its positive tone: members of a local mosque in the city of York had engaged with right-wing marchers with offers of tea and a friendly game of football [26], [27]. Many U.K. media outlets covered the story² and it was reiterated by commentators seeking to calm tensions. Our analysis of the data collected by Sentinel indicates that a very few far-right supporters actually responded to the calls on Twitter to march in York, possibly because a major national protest had taken place elsewhere in the U.K. the day before. Going further, the study in [25] draws on the data collected and analyzed via Sentinel to present a model of social reaction in the aftermath of terrorist attacks. A subsequent study [27] validated a model of social conflict using the data set and Sentinel tools.

4) *Discussion*: One key issue that emerges from this paper is the need to obtain ground truth for events reported on social media, such as the York Mosque incident. There is no doubt, however, that Twitter has emerged as a key tool for mobilizing support for issues and conducting campaigns. As we saw also with the escalation of the “poor policing” story in relation to the pro-Gaza protest march in Cardiff in July 2014, Twitter in conjunction with other sources such as YouTube is not merely a side channel for carrying information about real-world events, but has become a legitimate space in which campaigns are conducted in the virtual world.

C. Study 3: Assessing Community Impacts of a Major Event

1) *Purpose*: In September 2014, the South Wales region hosted the international NATO Summit, involving leaders and senior delegates from around 60 countries. The event was described as the largest ever peacetime security operation in the U.K., with over 9000 police officers assigned to the event, and significant disruption caused to the local community by the preparations and Summit itself. The Summit provided an opportunity to apply Sentinel to real-time study of a major planned event in a city region over a three-month period.

2) *Method*: Similar to the second experiment above, the basis for data collection for the NATO Summit exercise was an existing channel, in this case the South Wales one. Here, the event-specific channel was extended to include a set of terms relating to NATO, the Summit, its venues, scheduled protests, and groups both pro- and anti- the event. Being local to the area, the team had the opportunity to conduct

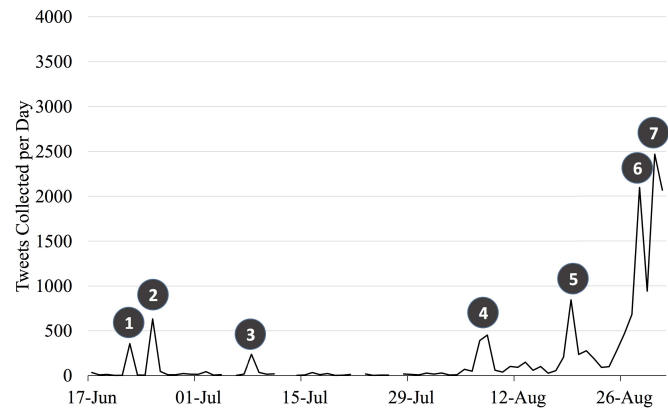


Fig. 14. Key events prior to Summit identified by tweet volume for terms containing “Summit.”

a natural experiment, running this exercise as a combination of “hackathon” and ethnographic study. The channel began collecting data 90 days prior to the event (in early July 2014) and the computer science team created Sentinel apps, including SentiSum and SentiNow (see Section IV) to support this paper. During the week of the Summit (September 1–5, 2014), a team of eight analysts monitored the unfolding situation both in the lab, using the Sentinel tools in an “eyes-on” capacity while also monitoring mainstream and social media manually, and in the field. A key focus of this paper was to use the Sentinel tools to task field teams to obtain eyewitness information (ground truth) on events signaled via Sentinel from Twitter data.

3) *Results*: Data volumes from the NATO Summit channel over a hundred days from July to mid-September 2014 are shown in Fig. 14. By this stage, the team had become adept at tuning the channel with additional topical and spatial terms and Twitter accounts of interest. The screenshot in Fig. 9 (see Section IV) is from this study, showing the situation at noon on the first day of the Summit (September 4) while a protest march was in progress on Chepstow Road in Newport, South Wales. The SentiSum screenshot in Fig. 10 is also from this paper, showing tweet volumes and sentiment profiles for tweets relating to the Summit for the month prior to the event and the Summit itself.

Each peak in Fig. 14 corresponds to significant online reaction to some aspect of the buildup to the Summit: 1) first major U.K. national news coverage of the Summit; 2) revealing of the Summit logo, with some local dissatisfaction over the choice of symbols included; 3) announcement that many local schools would be closed around the time of the Summit; 4) announcement of local road closures and a no-fly zone during the Summit; 5) installation of security fencing around Summit venues in Cardiff city centre, which became known as the “ring of steel,” causing significant traffic disruption; 6) the “ring of steel” being reported on BBC U.K. national news; and 7) first protest march in Newport opposing the Summit. The coloring of these peaks (see Fig. 10) shows significant negative reaction to each event. SentiSum visualizations of the ongoing online sentiment toward the Summit between June and mid-September (the week after the event) revealed a predominantly

²For example: <http://www.theguardian.com/uk/2013/may/27/york-mosque-protest-tea-biscuits>.

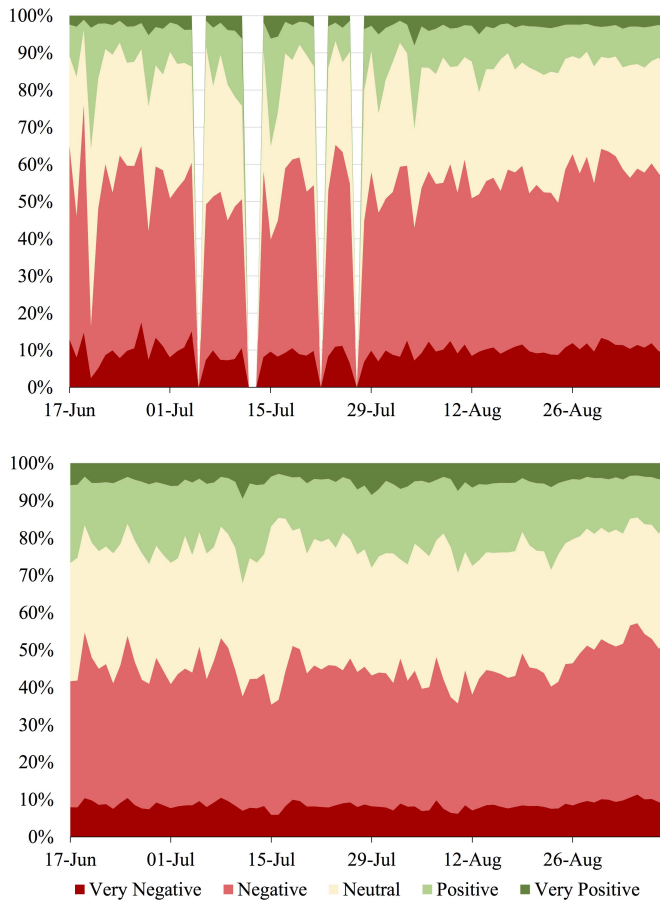


Fig. 15. Sentiment of Summit-related terms (top) versus all terms (bottom), June–August 2014.

negative view: Fig. 15 shows that the sentiment profile for signal terms relating to “Summit” was generally about 10% more negative than the baseline defined as the profile for all signal terms generated by FlexiTerm for the channel.

Once again, Sentinel proved effective at detecting and localizing key incidents in relation to a major event. What was new here was the tasking of field teams for ground-truthing events signaled from social media. Of particular interest were two questions in relation to protests: 1) how many protesters were present and 2) what was their mood? The former is typically a key area of uncertainty in relation to such events, with supporters tending to overestimate the number of participants, and opponents tending to underestimate. In some cases, calls to protest at a particular place and time were signaled on Twitter where our field teams confirmed that nobody had actually participated. The second question, that of the mood of the participants, is a case where independent human observers are a much more effective sensor than social media, where detection of sentiment in text is acknowledged as a hard problem [28]. This paper was one of the key technology transition outputs of a ten-year, 25-partner U.S./U.K. interdisciplinary research program in network and information science [29].

4) *Discussion*: In relation to the Summit and associated events, the key finding from this paper was that Sentinel was very effective at providing situation awareness of impacts

on the local community in terms of disruption (especially transport, security, and protests) and positive and negative public reactions. The value of Twitter as a real-time sensor is underlined by our experiences here. The closed-loop tasking of field teams to obtain ground truth on events detected from Twitter data also proved effective and valuable in situational understanding. More generally, this paper showed a high degree of reusability of elements of the Sentinel infrastructure, building on experiences in the first two experiments. The Summit channel was an expanded and slightly repurposed variant of our ongoing South Wales collection, with additional spatial, topical, and actor parameters. Some of these terms were then incorporated back into the original channel for ongoing use. Similarly, the ontology was expanded with additional concepts, and the SentiSum and SentiNow apps were useful beyond the Summit experiment.

VI. RELATED WORK

This paper is based on the principle that social media, in general, and Twitter, in particular, serves as a human-based sensor network [7], [30] in which people are data sources, rather than technology-based sensing devices. Moreover, we take the view that social media should be considered as one of multiple sources in attempting to make sense of some situation [9]. Our social science team has for many years conducted interviews with community members to gain a deep understanding of issues relating to crime and social disorder [17], [18]; Sentinel was originally conceived to offer a fast-time complement to these traditional interviewing methods.

A recent survey [10] identifies several key requirements in delivering semantic-based information processing of social media streams, including meaningful visualizations (entity-based, sentiment-based, and time-based) at different granularities, in real-time, supporting user interactivity, with integrated search, and scalable processing. We concur with all of these, and have addressed them in the design of Sentinel. In this sense, Sentinel is comparable with other systems, including Media Watch [31], Tweetgeist [32], TwitInfo [33], Twitris [34], and GATE [35], though none of them have our specific focus of supporting social science research via the incorporation of bespoke social science models. Moreover, we would add *openness* and *codesign* as two additional requirements, emphasizing the importance of a user-led approach to creating analytic tools supported by an open system architecture.

Use of social media for human-based sensing and situational understanding is complementary to the various crowdsourcing approaches that have emerged in recent years ([12] provides a recent survey), several of which focus on crisis response and disaster relief; e.g., Bellingcat (www.bellingcat.com), LRA Crisis Tracker (www.lracrisistracker.com), and Ushahidi (www.ushahidi.com). The difference is that, in the crowdsourcing approaches, people are directed to contribute pieces of information relating to some situation or query whereas, in the social media approaches, postings are spontaneous and triggered by external events.

In Section II, we characterized two modes of use for the Sentinel platform: bottom-up (data-driven) and

top-down (ontology-driven). This characterization is related to the keyword-based versus topic-based distinction made in [11] in that our channels for bottom-up processing are largely (though not exclusively) defined in terms of keyword sets, and the ontology-driven top-down framing is topic-based. In our senses of the terms bottom-up and top-down, related systems tend to characterize themselves as predominantly one or the other whereas, in creating Sentinel, we have endeavored to support both modes equally, in order for the platform to be maximally open to a variety of social science research uses. We would argue that Sentinel is, therefore, multifaceted in terms of [11].

Considering systems that provide top-down analyses, the Twitcident system focuses on supporting user-driven search and filtering on social media in relation to events identified in feeds from the emergency services [36]. Social Sensor [37] takes a top-down approach in that it is configurable to track social media traffic around a specific event, and in this respect it is topic-based in terms of [11], even though it also has a bottom-up element in terms of supporting newsfeed monitoring.

In terms of bottom-up analyses, there is considerable work in event detection using social media streams, particularly Twitter. Sakaki *et al.* [5] built a probabilistic spatiotemporal model to perform event detection using tweets in relation to earthquakes. Work by Vavliakis *et al.* [38] integrates named entity recognition (NER), topic discovery and clustering, and peak detection techniques to identify events in streamed social media data. The ReDites system [39] builds upon prior work in topic detection and tracking [40] with a focus on improving precision over recall. MEMAS demonstrated the effectiveness of social media as a sensor for detecting large-scale events via bottom-up analysis, even though showed that relatively local events often do not generate enough social media signal to be detectable via bottom-up analysis [41]. While we have had some success in detecting local events using Sentinel, this is usually due to our semantic channel-based approach tending to select data on local search terms, amplifying small signals, the “Mill Lane” example in Study 1 being a case in point.

This paper is not focused on event detection specifically but is broadly compatible with the above work. Specific event detection algorithms—[42] presents a recent comparative analysis—could in principle be plugged-into the Sentinel pipeline, to be applied concurrently with our ATR approach and the results potentially joined with the signal terms produced by FlexiTerm. As we have seen in the pilot studies, the generated terms often relate to events, especially when discussion of off-line events dominates online social media activity on a channel. However, our signal terms are more general in nature, picking up common topics of online conversation and, being noun phrases, tend to be more semantically meaningful and contextualized than keywords (e.g., “armed police” instead of “armed” or “police”).

A significant issue in using social media as a sensor network is the veracity of information obtainable from such open channels. It is a well-known issue that misinformation flows as well, if not even better than actual information on social media [43]. Sometimes this is due to malice, attempts to be

humorous, or propaganda as we observed this Section V-B regarding the “York Mosque tea party” story. This issue has been characterized as a reliable sensing problem in [44] and more broadly in [7]. The authors model human participants as sources of unknown reliability generating data of uncertain provenance, and show that an estimation-theoretic problem can be used to optimize filtering of correct observations. Such an approach could be incorporated into Sentinel in the future; at this stage, however, the tool is often used in an exploratory mode and we are as interested in the flow of misinformation, rumour, and propaganda as we are in “truth” [25].

VII. CONCLUSION AND DISCUSSION

Experience gained from the pilot studies (see Section V) indicates that Sentinel fulfilled our objectives set out in Section I to create an open platform that allows social and computing scientists to codesign useful analytic components and apps, able to semantically enrich social media data in both a bottom-up and top-down manner. The pilots exemplified two types of study: those focused on a specific geospatial region and those focused on a particular topic. Some of these are ongoing while others were focused on a specific bounded time period.

In these projects, we have observed two main kinds of use of Sentinel: 1) “eyes-on,” where a user has a Sentinel app on their device, possibly in conjunction with other Twitter and social media apps, typically while some specific situation is ongoing and 2) where some external channel (e.g., a news media feed or a personal message) indicates to a user that something is happening or has happened, and the user then goes into Sentinel to explore what the platform has picked up on that situation. Type I activity tends to be dominated by bottom-up uses of Sentinel, including scanning the signal feed, viewing the photo collage (which often provides an informative visual overview of some situation when several people are tweeting images from a scene), or accessing geotagged tweets from an area of interest. Type II activity tends to be driven by the top-down functionalities, including text and ontology-based searches, and using the timeline to access signals from a specific period. However, both kinds of use typically involve a combination of bottom-up and top-down activity. For example, an interesting image spotted in “eyes-on” mode will commonly lead to top-down searches for related activity.

One of the most important insights we gleaned from the pilots has been the importance of campaigning behavior in the social media space. In all the studies, we observed interesting parallels between physical and online campaigning. In the case of the July 2014 Cardiff protest, we saw how a relatively minor incident of violence was in effect magnified by social media activity on YouTube and Twitter, resulting in greater coverage for the protest in mainstream media. In the Woolwich murder case, we saw how the York Mosque narrative of peaceful response was very effectively broadcast, regardless of the small number of right-wing protesters actually engaged. In the NATO Summit case, we saw repeated examples of online calls to protest which were answered by few or no people physically appearing. These cases raise interesting sociological questions

of whether online protesting is in some circumstances replacing physical protests, and whether a protest can be said to have occurred if it only happened online?

In terms of immediate future work, we are developing a notification system for Sentinel. We deliberately postponed the design of such a system until we had gained a good amount of experience in using the platform, in order to better tailor the notifications to items of significant likely interest. In relation to this, we also intend to revisit the topic of event detection and tracking, drawing on some of the works covered in Section VI.

We also plan to incorporate methods for expanding the ontology dynamically, using the FlexiTerm-generated signal terms in conjunction with NER. We are also experimenting with methods for incorporating user-defined annotations as a complement to the ontology for top-down analysis. We will also enhance Sentinel's ontology indexing to perform NER using the NER tool provided by the StanfordNLP toolkit (nlp.stanford.edu). The current Ontology Indexer (see Section III) will be replaced with the PathNER tool (www.biomedcentral.com), which provides a convenient mechanism for defining an entity set that can be mapped to texts. The set will consist of the Sentinel Ontology elements, including synonyms. Results from the Stanford NER tool (identifying people, places, and organizations) will be added to the Sentinel Ontology matches. This enhancement of the Ontology Indexer module will be performed in concordance with the development of a text classification component of the Sentinel pipeline, where the ontology and entity matches are to be offered up as potential feature vectors in the classification exercise.

In the work to date, we have focused upon evaluating the Sentinel architecture and tools via a series of case studies, in the context of social science research as exemplified by the three studies described in Section V, rather than via user studies aimed at measuring task performance ([45] provides a good example of such a study for visualization approaches). However, this paper involves performing a user study of the top-down Sentinel interface on performing ontology-based search queries using the tool.

As a final comment, this paper has focused on improving situational understanding using social media as a sensor. Increasingly, use of social media in public services, including policing, is seen as not only a (one-way) sensor but as a means of two-way engagement. Examples include Staffordshire Police in the U.K. using Facebook for community engagement (policemediablog.com), and the Spanish Police's use of Twitter to crowdsource eyewitness reports.³ In fact, following the posting of YouTube videos of the violent incident reported in Section V, Study 1, South Wales police issued calls on social media for people to help identify individuals in the videos. We see considerable potential for tools like Sentinel to support both sensing and effecting tasks in the future, including both crowdsourcing and conversational interactions with people on social media. This is among the findings of a

recently concluded study commissioned by the U.K. College of Policing [46].

ACKNOWLEDGMENT

The authors thank Prof. C. Jones for advice on geospatial aspects of the ontology. They thank D. Braines of IBM Emerging Technology, Hursley Park, U.K., for valuable assistance with the NATO Summit case study and input on policing uses of social media.

REFERENCES

- [1] H. Gao, G. Barbier, and R. Goolsby, "Harnessing the crowdsourcing power of social media for disaster relief," *IEEE Intell. Syst.*, vol. 26, no. 3, pp. 10–14, May/Jun. 2011.
- [2] D. Lazer *et al.*, "Computational social science," *Science*, vol. 323, no. 5915, pp. 721–723, 2009.
- [3] D. Wang, L. Kaplan, H. Le, and T. Abdelzaher, "On truth discovery in social sensing: A maximum likelihood estimation approach," in *Proc. 11th Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, Apr. 2012, pp. 233–244.
- [4] M. Osborne and M. Dredze, "Facebook, Twitter and Google plus for breaking news: Is there a winner?" in *Proc. AAAI Conf. Web Social Media*, 2014, pp. 611–614.
- [5] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes Twitter users: Real-time event detection by social sensors," in *Proc. 19th Int. Conf. World Wide Web*, 2010, pp. 851–860.
- [6] D. Gayo-Avello, "A meta-analysis of state-of-the-art electoral prediction from Twitter data," *Social Sci. Comput. Rev.*, vol. 31, no. 6, pp. 649–679, 2013.
- [7] D. Wang, T. Abdelzaher, and L. Kaplan, *Social Sensing: Building Reliable Systems on Unreliable Data*. San Mateo, CA, USA: Morgan Kaufmann, 2015.
- [8] D. L. Hall and J. M. Jordan, *Human-Centered Information Fusion*. Norwood, MA, USA: Artech House, 2010.
- [9] P. Pirolli and S. Card, "The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis," in *Proc. Int. Conf. Intell. Anal.*, 2005, pp. 1–6.
- [10] K. Bontcheva and D. Rout, "Making sense of social media streams through semantics: A survey," *Semantic Web*, vol. 5, no. 5, pp. 373–403, 2014.
- [11] Y. Wu, N. Cao, D. Gotz, Y.-P. Tan, and D. A. Keim, "A survey on visual analytics of social media data," *IEEE Trans. Multimedia*, vol. 18, no. 11, pp. 2135–2148, Nov. 2016.
- [12] M. Imran, C. Castillo, F. Diaz, and S. Vieweg, "Processing social media messages in mass emergency: A survey," *ACM Comput. Surv.*, vol. 47, no. 4, 2015, Art. no. 67.
- [13] I. Spasić, M. Greenwood, A. Preece, N. Francis, and G. Elwyn, "FlexiTerm: A flexible term recognition method," *J. Biomed. Semantics*, vol. 4, p. 27, Oct. 2013.
- [14] M. Uschold, "Building ontologies: Towards a unified methodology," in *Proc. 16th Annu. Conf. Brit. Comput. Soc. Spec. Group Expert Syst.*, 1996, pp. 1–20.
- [15] R. B. Altman, M. Buda, X. J. Chai, M. W. Carillo, R. O. Chen, and N. F. Abernethy, "RiboWeb: An ontology-based system for collaborative molecular biology," *IEEE Intell. Syst. Appl.*, vol. 14, no. 5, pp. 68–76, Sep. 1999.
- [16] P. Grenon, B. Smith, and L. Goldberg, "Biodynamic ontology: Applying BFO in the biomedical domain," in *Ontologies in Medicine*. Amsterdam, The Netherlands: IOS Press, 2004, pp. 20–38.
- [17] M. Innes, *Signal Crimes: Social Reactions to Crime, Disorder, and Control*. New York, NY, USA: Oxford, 2014.
- [18] T. Lowe and M. Innes, "Can we speak in confidence? community intelligence and neighbourhood policing v2.0," *Policing Soc., Int. J. Res. Policy*, vol. 22, no. 3, pp. 295–316, 2012.
- [19] J. Day-Richter, M. A. Harris, M. Haendel, and S. Lewis, "OBO-Edit—An ontology editor for biologists," *Bioinformatics*, vol. 23, no. 16, pp. 2198–2200, 2007.
- [20] World Wide Web Consortium. *OWL 2 Web Ontology Language Document Overview*, 2nd ed. World Wide Web Consortium (W3C), Dec. 2012. [Online]. Available: <http://www.w3.org/TR/owl2-overview/>
- [21] S. Vinoski, "Advanced message queuing protocol," *IEEE Internet Comput.*, vol. 10, no. 6, pp. 87–89, Nov./Dec. 2006.

³"Police in Spain Arm Themselves With Social Media to Fight Crime": www.nytimes.com/2014/09/08/business/media/.

- [22] D. Rogers *et al.*, “Bundle and pool architecture for multi-language, robust, scalable workflow executions,” *J. Grid Comput.*, vol. 11, no. 3, pp. 457–480, 2013.
- [23] R. Socher *et al.*, “Recursive deep models for semantic compositionality over a sentiment treebank,” in *Proc Empirical Methods Natural Lang. Process. (EMNLP)*, 2013, pp. 1631–1642.
- [24] J. Mahmud, J. Nichols, and C. Drews, “Home location identification of Twitter users,” *ACM Trans. Intell. Syst. Technol.*, vol. 5, no. 3, p. 47, 2014.
- [25] M. Innes, C. Roberts, A. Preece, and D. Rogers, “Ten ‘Rs’ of social reaction: Using social media to analyse the ‘post-event’ impacts of the murder of Lee Rigby,” *Terrorism Political Violence*, Jul. 2016. [Online]. Available: <https://doi.org/10.1080/09546553.2016.1180289>
- [26] C. Roberts, M. Innes, A. Preece, and I. Spasić, “Soft facts and spontaneous community mobilisation: The role of rumour after major crime events,” in *Data for Good: How Big Open Data can be Used for Common Good*, P. Baecck, Ed. London, U.K.: Nesta, 2015, pp. 37–43.
- [27] C. Roberts, M. Innes, A. Preece, and D. Rogers, “After woolwich: Analyzing open source communications to understand the interactive and multi-polar dynamics of the arc of conflict,” *Brit. J. Criminol.*, Apr. 2017. [Online]. Available: <https://doi.org/10.1093/bjc/azx024>
- [28] B. Liu, *Sentiment Analysis and Opinion Mining*. San Rafael, CA, USA: Morgan & Claypool, 2012.
- [29] G. Bent *et al.*, *Network and Information Sciences International Technology Alliance*. New York City, NY, USA: IBM, 2016.
- [30] M. Srivastava, T. Abdelzaher, and B. Szymanski, “Human-centric sensing,” *Philos. Trans. Roy. Soc. London A, Math. Phys. Sci.*, vol. 370, no. 1958, pp. 176–197, 2012.
- [31] A. Hubmann-Haidvogel, A. M. Brasoveanu, A. Scharl, M. Sabou, and S. Gindl, “Visualizing contextual and dynamic features of micropost streams,” in *Proc. MSM Workshop*, vol. 838, 2012, pp. 35–40.
- [32] D. A. Shamma, L. Kennedy, and E. F. Churchill, “Tweetgeist: Can the twitter timeline reveal the structure of broadcast events?” in *Proc. CSCW*, 2010.
- [33] A. Marcus, M. S. Bernstein, O. Badar, D. R. Karger, S. Madden, and R. C. Miller, “Twitinfo: Aggregating and visualizing microblogs for event exploration,” in *Proc. Conf. Hum. Factors Comput. Syst. (CHI)*, 2011, pp. 227–236.
- [34] M. Nagarajan, K. Gomadam, A. P. Sheth, A. Ranabahu, R. Mutharaju, and A. Jadhav, “Spatio-temporal thematic analysis of citizen sensor data: Challenges and experiences,” in *Proc Int. Conf. Web Inf. Syst. Eng.*, 2009, pp. 539–553.
- [35] D. Maynard, I. Roberts, M. A. Greenwood, D. Rout, and K. Bontcheva, “A framework for real-time semantic social media analysis,” *Web Semantics, Sci., Services Agents World Wide Web*, vol. 44, pp. 75–88, May 2017.
- [36] F. Abel, C. Hauff, G.-J. Houben, R. Stronkman, and K. Tao, “Semantics + filtering + search = Twitcident. Exploring information in social Web streams,” in *Proc. 23rd ACM Conf. Hypertext Social Media*, 2012, pp. 285–294.
- [37] L. M. Aiello *et al.*, “Sensing trending topics in Twitter,” *IEEE Trans. Multimedia*, vol. 15, no. 6, pp. 1268–1282, Oct. 2013.
- [38] K. N. Vavliakis, A. L. Symeonidis, and P. A. Mitkas, “Event identification in Web social media through named entity recognition and topic modeling,” *Data Knowl. Eng.*, vol. 88, pp. 1–24, Nov. 2013.
- [39] M. Osborne *et al.*, “Real-time detection, tracking, and monitoring of automatically discovered events in social media,” in *Proc. 52nd Annu. Meeting Assoc. Comput. Linguistics*, 2014, pp. 37–42.
- [40] S. Petrović, M. Osborne, and V. Lavrenko, “Streaming first story detection with application to Twitter,” in *Proc. Hum. Lang. Technol., Annu. Conf. North Amer. Chapter Assoc. Comput. Linguistics*, 2010, pp. 181–189.
- [41] E. Bahir and A. Peled, “Real-time major events monitoring and alert system through social networks,” *J. Contingencies Crisis Manage.*, vol. 23, no. 4, pp. 210–220, 2015.
- [42] A. Weiler, M. Grossniklaus, and M. H. Scholl, “An evaluation of the run-time and task-based performance of event detection techniques for Twitter,” *Inf. Syst.*, vol. 62, pp. 207–219, Dec. 2016.
- [43] F. Jin *et al.*, “Misinformation propagation in the age of Twitter,” *Computer*, vol. 47, no. 12, pp. 90–94, Dec. 2014.
- [44] D. Wang *et al.*, “Using humans as sensors: An estimation-theoretic perspective,” in *Proc. 13th Int. Symp. Inf. Process. Sensor Netw. (IPSN)*, Apr. 2014, pp. 35–46.
- [45] A. Weiler, M. Grossniklaus, and M. H. Scholl, “Situation monitoring of urban areas using social media data streams,” *Inf. Syst.*, vol. 57, pp. 129–141, Apr. 2016.
- [46] M. Innes and C. Roberts, Eds., “Open source communications analytics research (OSCAR) development centre,” Univ. Police Sci. Inst., Cardiff, U.K., Final Rep., 2017.

Alun Preece received the Ph.D. degree in computer science from the University of Wales, Swansea, U.K., in 1989.

He is the Co-Director of the Crime and Security Research Institute, Cardiff University, Cardiff, U.K. He is currently the U.K. Academic Technical Area Leader for the Distributed Analytics and Information Sciences International Technology Alliance funded by the U.K. Ministry of Defence and the U.S. Army Research Laboratory. His current research interests include decision support and human-computer collaboration.

Irena Spasić received the Ph.D. degree in computer science from the University of Salford, Salford, U.K., in 2004.

She is a fellow of the Crime and Security Research Institute, Cardiff University, Cardiff, U.K. Following posts at the Universities of Belgrade, Salford, and Manchester, she joined the Cardiff School of Computer Science and Informatics, Cardiff, in 2010, and became a Full Professor in 2016. Her current research interests include text mining, knowledge representation, and machine learning.

Dr. Spasić co-founded the U.K. Healthcare Text Analytics Research Network.

Kieran Evans received the B.Sc. degree in computer science from Cardiff University, Cardiff, U.K.

He was a Researcher in computer science on EU and Wellcome Trust Projects. He has been a researcher with the Crime and Security Research Institute, Cardiff University. He is currently the Director of Development at ASB Consulting LLC, Tysons, VA, USA.

David Rogers is currently pursuing the Ph.D. degree with the School of Computer Science and Informatics, Cardiff University, Cardiff, U.K. His Ph.D. thesis was titled Text Mining of Extremist Narratives on the Web.

He is currently a Research Assistant with Cardiff University. He was involved in a number of EC-funded projects focused on supporting heterogeneous workflow application infrastructure. He is now applying his interest in Big Data and the Semantic Web to his current interdisciplinary work within the Crime and Security Research Institute.

William Webberley received the B.Sc. degree (Hons.) in computer science and the Ph.D. degree in solving the unfiltered feed problem experienced in online social networks from Cardiff University, Cardiff, U.K., in 2010 and 2015, respectively.

Since 2015, he has been a Researcher with the Crime and Security Research Institute, Cardiff University, with a focus on distributed and decentralized agent-agent and human-agent systems for knowledge base building and task assignment.

Colin Roberts is the Senior Research Fellow of the Universities Police Science Institute within the Crime and Security Research Institute, Cardiff University, Cardiff, U.K. He led the EC TARDIS Program developing social media analytics to detect and counter online radicalization, and was a Program Manager with the Open Source Communications Analytics Research Centre, Cardiff University. His current research interests include social media analytics of the time dynamics of conflict, multipolar dynamics of extremism, and intelligence processes.

Martin Innes is the Director of the Crime and Security Research Institute and Universities’ Police Science Institute, Cardiff University, Cardiff, U.K. He is the Author of the books *Signal Crimes* and *Investigating Murder* (Oxford University Press). His current research interests include counter-terrorism policing and applying social media analytics to the study of social reactions to terrorist attacks.